Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 76 (2015) 192 – 197

Procedia Computer Science

2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS 2015)

Tracking of Multiple Light Sources Using Computer Vision for Underwater Docking

M. F. Yahya[a]*, M. R. Arshad[a]

[a]*Underwater, Control and Robotics Group, School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Engineering Campus, 14300, Nibong Tebal, Seberang Perai Selatan, Pulau Pinang, Malaysia.*

Abstract

An underwater docking platform can be used as a recharging station for autonomous underwater vehicle (AUV) to prolong its operation. Prior researches indicate that optical sensor is used by an AUV to capture docking station images so as to guide the vehicle towards the station. However, docking using vision is difficult to realize due to the problem of determining the position and orientation of the station with respect to the AUV. Consequently, a tracking system using computer vision is proposed in this paper. Obtained results showed that the tracking method was able to recognize the light sources placed on the docking station which is represented by artificial targets with True Positive Rate of more than 0.9. Furthermore, it was found from this study that as the camera approaching speed towards the target increases, successful recognition decreases. Therefore, an AUV have to propel slowly and steadily to have a successful docking operation.

Peer-review under responsibility of organizing committee of the 2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS 2015)

Keywords: tracking; underwater docking; computer vision; light sources

1. Introduction

Ocean is undoubtedly an unfamiliar territory where valuable resources lie within. Where the physicality of human cannot venture further under the ocean, autonomous underwater vehicle (AUV) can perform the exploration without human supervision. Additionally, some of the common applications of AUV include seabed mapping, pipe inspection, and ocean study. Although most of AUV can remain unattended by human operator, it comes with a trade-off of having limited power supply. Even more, its design and processing is bounded by how long the battery can last for a single operation. After the battery has exhausted, the AUV has to be retrieved and recharged from time to time which makes the process tedious and costly. Towards this end, having an underwater docking system is one of the alternatives to solve this problem. Apart from being a station for the AUV to recharge its battery, it can be used as a launch and recovery of underwater vehicles and allows the AUV to perform station-keeping.

While acoustic sensor remains to be the most used sensor for many underwater applications, vision is more suited for close range application per se docking. Related research for AUV docking had been conducted for torpedo shaped AUV [1] and box shaped AUV [2]. A camera is used to guide the torpedo shaped AUV towards a circular shaped entrance docking station during homing. On the other hand, acoustic and vision sensors are combined to achieve 50% of docking success trials for box-shaped AUV.

* Corresponding author. Tel.: +60137871657
E-mail address: mohdfaid85@gmail.com

Peer-review under responsibility of organizing committee of the 2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS 2015)
doi:10.1016/j.procs.2015.12.339

One of the most important aspects of successful underwater docking operation is the ability to continuously track the location of the station from captured video frames. Computer vision had been used to process video frames and identify light sources installed on the station [1]. Apart for docking, underwater tracking system had mostly been developed for pipeline [3] and cable inspections [4]. Before the target can be tracked, it needs to be detected and recognized. Several works had been conducted for underwater target detection and recognition. In one work, feature-based approach and template-based approach had been studied to detect underwater objects [5]. Additionally, optical flow and mean shift tracking methods are compared in the same study. In another work, an acoustic camera is used to perform underwater object recognition [6]. However, acoustic camera requires extensive processing and the captured image quality is degraded tremendously when compared to using optical sensor.

For underwater docking, techniques such as background subtraction [7] and codebook [8] are not applicable since the camera attached on the AUV is in motion together with the AUV. Those techniques only suitable if the background scenery is fixed. There are some works where tracking is realizable through a moving camera. Kalman filter based on the center of gravity of an object region method was presented to track object while a camera is in motion [9]. Then, there is also tracking objects by using a moving binocular camera [10]. Although the method can track 3 dimensional objects, processing two video frames simultaneously and continuously will increase the processing time and therefore makes the frame rate slow. In other work, fusion of texture and color are used to segment object of interest from an image [11]. Although texture is important to obtain specific features of the desired tracked object, color is much more important to distinguish the object of interest in an image.

In this paper, tracking of light sources by using color information from a single moving camera had been proposed. The study includes real-time vision of image processing, detection, and recognition of the targets. The main focus is to find a relation between varying the camera approaching speed towards the target to the success of the target recognition. The contribution of this work includes a novel algorithm to recognize the target and evaluating the proposed algorithm in real time.

2. Tracking Method

The work proposed in this paper is on development of a method to track light sources originating from an underwater docking station. The method consists of image acquisition, color space conversion, color thresholding, morphological operations, target detection, and recognition. The flow of the proposed method is shown in Fig. 1. Accordingly, the image after each process is shown in Fig. 2. Initially, an image is acquired from a charged-coupled device (CCD) image sensor in the form of a webcam. The webcam is suitable for real time application since it produces frame rate of at least 15 Hz at a resolution of 680 pixels wide to 480 pixels height. Fig. 2 (a) shows an image acquired from the webcam where there is a light source as well as an artificial object.

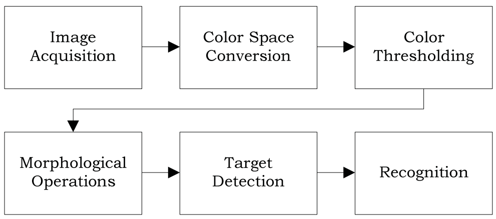

Fig. 1. Tracking method overview.

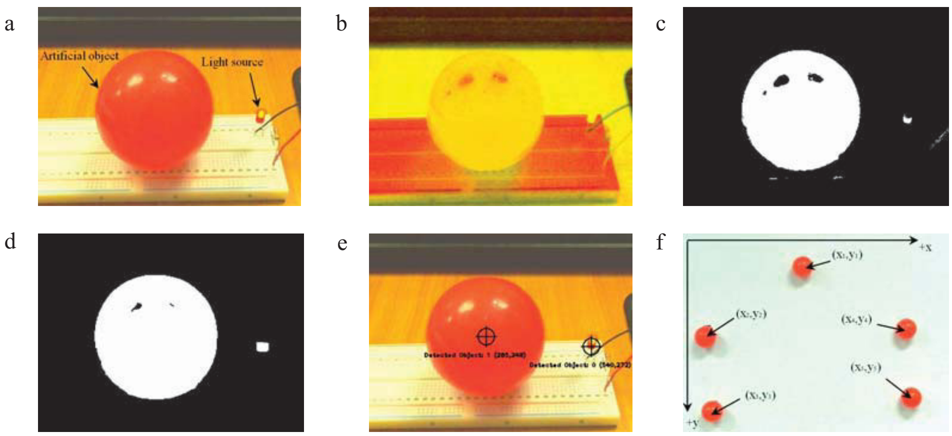

Fig. 2. Images from each step in the tracking method: (a) original (b) HSV (c) color thresholding (d) morphological opening (e) detected targets and (f) centroid for 5 targets and positive axes of x and y for recognition.

The image obtained is in Red-Green-Blue (RGB) color space. The components in RGB are correlated to the amount of light wavelengths hitting the objects in the captured image. Due to this, the discrimination of the object of interest from the surrounding by using RGB channel is difficult. Therefore, it needs to be converted to other color format for further processing. One of the alternatives is by using Hue-Saturation-Value (HSV) color space. Unlike RGB, HSV separates image intensity from color information for easier object of interest filtering. Fig. 2 (b) shows a converted HSV image from previous RGB image using the following formula

$$V \leftarrow \max(R,G,B) \tag{1}$$

$$S \leftarrow \begin{cases} \dfrac{V - \min(R,G,B)}{V} & \text{if } V \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$H \leftarrow \begin{cases} \dfrac{60(G-B)}{V-\min(R,G,B)} & \text{if } V = R \\ \dfrac{120+60(B-R)}{V-\min(R,G,B)} & \text{if } V = G \\ \dfrac{240+60(R-G)}{V-\min(R,G,B)} & \text{if } V = B \end{cases} \tag{3}$$

What follows is color thresholding. Color thresholding is one of the methods in image segmentation. In this proposed method, image segmentation is required to separate the light source from the surrounding. The produced result is a binary image consisting of black pixel of value 0 and white pixel of value 255. The selected values are 0 to 15 for H component, 150 to 255 for S component, and 246 to 255 for V component. Any values that fall within this specified range are converted to white pixels and values outside of the range are converted to black pixels. Fig. 2 (c) shows an image with applied HSV color filter. It is important to note that the artificial object is selected as a substitute to light source since they both can be thresholded with the same values of HSV color filter as depicted in the same Fig. 2 (c).

The image obtained after color thresholding contains noisy white pixels. Therefore, morphological opening is used to remove these noises while preserving the shape and size of the target [12]. The process involved twice erosion followed by twice dilation. Erosion eliminates small noises and dilation reconstructs the target back to its original shape. A structuring element of rectangular 3 by 3 was used for erosion and a rectangular 8 by 8 for dilation. Fig. 2 (d) shows the effect of morphological opening.

In order to detect targets from the pre-processed image, there are 3 courses of action. The first is contour finding where each connected component of white pixels in the image is classified as separate objects. Then, the second is calculating moment (weighted average of pixels' intensity) for each contour using Green's theorem [13]. From the calculated moment, the area and centroid for each contour are obtained. The third is filtering the contours whereby minimum area is set to 15 by 15 pixels. Area of contours lower than this are considered to be noises. Fig. 2 (e) shows the detected object with labeled markings.

Detection data would not suffice to give the position and orientation of the entrance of the docking station. In this case, recognition is required to recognize all of the detected targets. This information is useful for the AUV to have a generated reference angle of attack during docking. There are 5 targets need to recognized which confirms and validates as a docking station. The recognition is based from evaluating positions of centroid of these detected targets. The centroids are illustrated in Fig. 2 (f) with shown x and y axes. The recognition rules are given as follow

Rule 1: $y_1 < \max(y_2, y_3, y_4, y_5)$ & $x_1 > \max(x_2, x_3)$ & $x_1 < \min(x_4, x_5)$

Rule 2: $x_2 < \min(x_1, x_4, x_5)$ & $y_2 > y_1$ & $y_2 < \min(y_3, y_5)$

Rule 3: $x_3 < \min(x_1, x_4, x_5)$ & $y_3 > \max(y_1, y_2, y_4)$

Rule 4: $x_4 > \max(x_1, x_2, x_3)$ & $y_4 > y_1$ & $y_4 < \min(y_3, y_5)$

Rule 5: $x_5 > \max(x_1, x_2, x_3)$ & $y_5 > \max(y_1, y_2, y_4)$

3. System Design and Experimental Setup

There are two platforms designed to perform the docking operation. The first platform is an AUV and the other platform is a docking station. Solidworks^{TM} software is used to design both of the platforms as shown in Fig. 3. Fig. 3 (a) shows the design of the AUV. A camera is installed inside the electronic container of the AUV to acquire underwater images. The camera is positioned at the frontal side of the container as depicted from the same Fig. 3. (a). Accordingly, an underwater docking station is designed with 5 light sources installed at the entrance of the docking station as shown in Fig. 3 (b). The frontal view of the docking station as well as positioning for each of the light source is shown in Fig. 4.

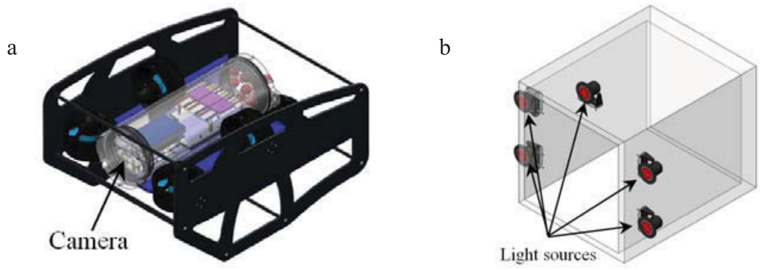

Fig. 3. Design of platforms: (a) AUV and camera position. (b) Docking station and light source installation.

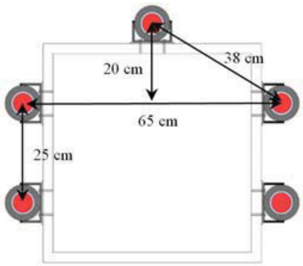

Fig. 4. Frontal view of docking station and light source positioning.

Note that in the current work, the designed AUV was represented by a single camera while light sources were represented by artificial targets. Although the test was conducted on ground and not in underwater, the concept remains the same with moving camera equivalent to AUV in motion and the tracked artificial targets equivalent to a docking station with light sources.

The experiments for the developed tracking method have been conducted at USM's Underwater Robotics Research Group (URRG) lab in Penang, Malaysia. The video streams are recorded using Logitech C920. For successful tracking, a camera is placed at 1.3 m from the floor to capture the image of the artificial targets and to obtain suitable Effective Detection Area (EDA). The experimental setup is as shown in Fig. 5. Apart from EDA, the other area shown in the same figure is Camera Blind Spot (CBS). CBS refers to the area where the camera is unable to capture surrounding area. CBS occurs as a result of limitation on focus area of camera lens. Note that the targets configuration is similar to that of light source positioning in Fig. 4 with the same measurement.

4. Experimental Results

Based from the experimental setup in Fig. 5, the camera started to record at about 6 meters away from the target, surged forward towards the target, and stopped at about 1 meter away from the target. The camera movement is controlled by an operator during recording. Fig. 6 shows captured images when the camera is at different distances.

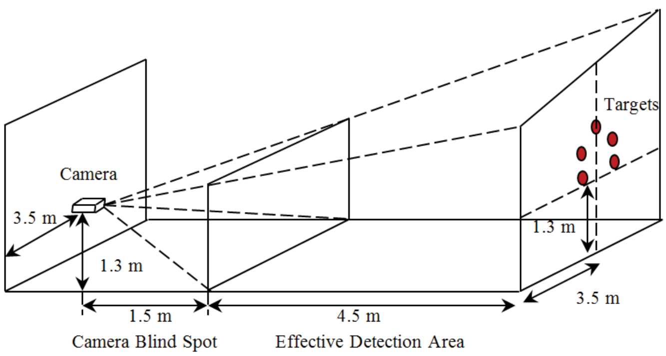

Fig. 5. Experimental setup.

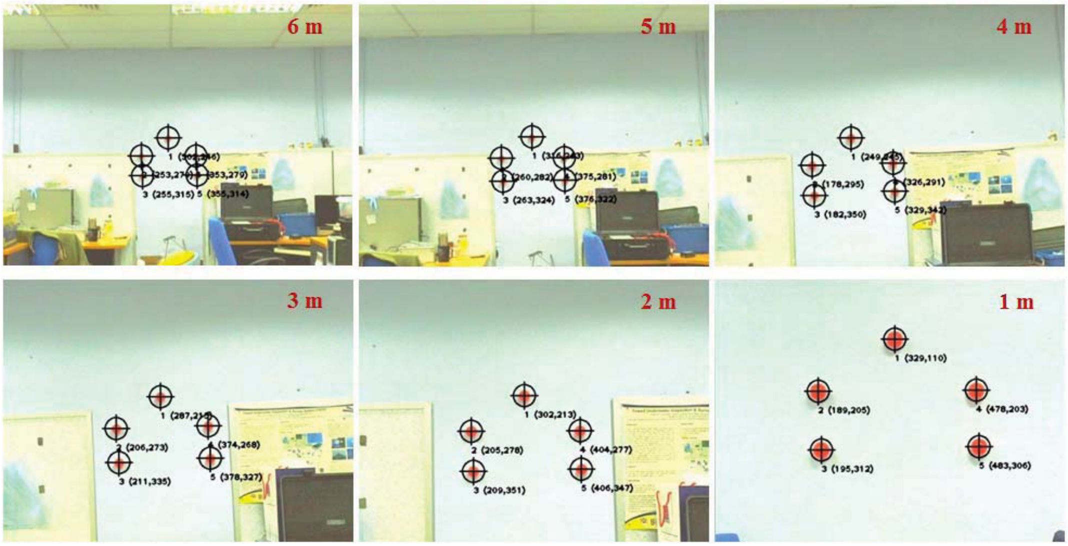

Fig. 6. Image captured at different distances.

For performance criteria, sensitivity or True Positive Rate (TPR) had been used. TPR is calculated from

$$TPR = \frac{TP}{TP + FN} \tag{4}$$

where TP is True Positive and FN is False Negative. In the context of this work, TP represents correct recognition of the 5 artificial targets and FN represents incorrectly unrecognized or in simpler term not being able to recognized the 5 artificial targets. Note that the test for False Positive (FP) and True Negative (TN) were not conducted since the focus was only on the 5 targets. Consequently, the accuracy of the image tracking is not calculated. For this test, high TRP value is desirable.

The effects of camera approach speed towards the target being recognized were investigated. The speed of the camera movement is set to ~0.1 m/s, ~0.15 m/s, and ~0.2 m/s. 5 trials were conducted for each set of speed so that average results can be calculated. The performance of the tracking method is shown in Table 1. Additionally, the table includes the total number of total frames (TF) and successful target detection (SD).

Table 1. Performance for tracking method.

Criteria	Average of 5 readings		
	~0.1 m/s	*~0.15 m/s*	*~0.2 m/s*
TF	692	506	388
SD	692	506	388
TP	667	482	352.2
FN	25.4	24	35.6
TPR	0.9636	0.9519	0.9089

The table showed that the proposed method was able to detect any available targets in every single video frame. This is probably due to the chosen color filter where the system easily distinguishes any target of interest from the surroundings. Furthermore, the speed at ~0.1 m/s showed the best results for average TPR value with 0.9636. The second best is 0.9519 for speed ~0.15 m/s and the third is 0.9089 for ~0.2 m/s. Based from the obtained result, the pattern shows that as the camera approaching speed increases, recognition accuracy of the target decreases.

At certain times, the tracker failed to track the target due to noises from the surrounding. Fig. 7 shows failure in tracking due to other object in the test field having the same color as the tracked targets. It is expected that in underwater environment, color noises from the surrounding will be much lesser but the quality of vision might be degraded and this may affect desirable distances to get the visual of the targets.

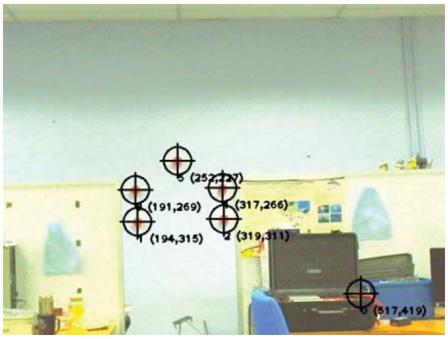

Fig. 7. Tracking failure.

5. Conclusion and Future Works

In this paper, target tracking with a moving camera for underwater docking using color have been studied. Basically, a tracking method is developed to keep track of artificial targets which represent the docking station light sources. The parameter to identify the light source in underwater environment is HSV color range. The results obtained were encouraging and ready to be applied in an AUV system. The camera speed in relation to recognition success shows that as the approaching speed increases recognition decreases. Thus, an AUV have to propel at a very slow and steady speed during docking operation so as to not lose track of the light sources and to guarantee docking success. Also, if the target is moving, it is expected that the vision algorithm is able to track the targets provided that the target moves within desirable speed which is subjected to further experimentation.

A few improvements to enhance the tracking method performance is by conducting a study of using different target color and implementing different target configurations to find the optimized setup. As for future work, the tracking method is going to be tested on an AUV to compare the tracking performance when in underwater and on ground.

Acknowledgement

This work was supported by Research University Incentive Grant (RUI) under account number 1001/PELECT/814234. (Project title: "Development of an Autonomous Underwater Vehicle (AUV) for Shallow Water Inspection and Monitoring with Novel Acoustic Tracking System").

References

1. J.-Y. Park, B.-h. Jun, P.-m. Lee, and J. Oh, "Experiments on vision guided docking of an autonomous underwater vehicle using one camera," Ocean Engineering, vol. 36, pp. 48-61, 1// 2009.
2. T. Maki, R. Shiroku, Y. Sato, T. Matsuda, T. Sakamaki, and T. Ura, "Docking method for hovering type AUVs by acoustic and visual positioning," in *Underwater Technology Symposium (UT), 2013 IEEE International*, 2013, pp. 1-6.
3. S. Krupinski, G. Allibert, H. Minh-Duc, and T. Hamel, "Pipeline tracking for fully-actuated autonomous underwater vehicle using visual servo control," in *American Control Conference (ACC), 2012*, 2012, pp. 6196-6202.
4. P. Drews, V. Kuhn, and S. Gomes, "Tracking System for Underwater Inspection Using Computer Vision," in *Offshore and Marine Technology: Science and Innovation (NAVTEC), 2012 International Conference on*, 2012, pp. 27-30.
5. D. Lee, G. Kim, D. Kim, H. Myung, and H.-T. Choi, "Vision-based object detection and tracking for autonomous navigation of underwater robots," *Ocean Engineering*, vol. 48, pp. 59-68, 7// 2012.
6. Y. Son-Cheol, K. Tae-Won, A. Asada, S. Weatherwax, B. Collins, and Y. Junku, "Development of High-Resolution Acoustic Camera based Real-Time Object Recognition System by using Autonomous Underwater Vehicles," in *OCEANS 2006*, 2006, pp. 1-6.
7. B. Karasulu and S. Korukoglu, "Moving object detection and tracking by using annealed background subtraction method in videos: Performance optimization," *Expert Systems with Applications*, vol. 39, pp. 33-43, 1// 2012.
8. Q. Zhong, Z. Qingqing, and G. Tengfei, "Moving Object Tracking Based on Codebook and Particle Filter," *Procedia Engineering*, vol. 29, pp. 174-178, // 2012.
9. W.-C. Hu, C.-H. Chen, T.-Y. Chen, D.-Y. Huang, and Z.-C. Wu, "Moving object detection and tracking from video captured by moving camera," *Journal of Visual Communication and Image Representation*, vol. 30, pp. 164-180, 7// 2015.
10. L. Cao, C. Wang, and J. Li, "Robust depth-based object tracking from a moving binocular camera," *Signal Processing*, vol. 112, pp. 154-161, 7// 2015.
11. E. Ozyildiz, N. Krahnstöver, and R. Sharma, "Adaptive texture and color segmentation for tracking moving objects," *Pattern Recognition*, vol. 35, pp. 2013-2029, 10// 2002.
12. E. R. Dougherty, "An Introduction to Morphological Image Processing", ISBN 0-8194-0845-X, 1992.
13. G. Green, "An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism," unpublished, pp. 10-12, 1828.